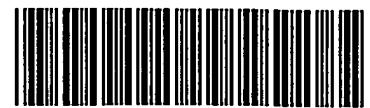# RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _09/938,703B_

Source: _____

Date Processed by STIC: _____

ENTERED

RAW SEQUENCE LISTING   DATE: 04/20/2005
PATENT APPLICATION: US/09/938,703B   TIME: 12:35:41

Input Set : **N:\Crf3\RULE60\09938703B.raw.txt**
Output Set: **N:\CRF4\04202005\I938703B.raw**

```
  1 <110> APPLICANT: Euroscreen S.A.
  2       SAMSON, Michel
  3       PARMENTIER, Marc
  4       VASSART, Gilbert
  5       LIBERT, Frederick
  6 <120> TITLE OF INVENTION: Methods for Identifying Compounds which Bind the Active CCR5
Chemokine
  7       Receptor
  8 <130> FILE REFERENCE: 9409/2023C
C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/938,703B
 10 <141> CURRENT FILING DATE: 2001-08-24
 11 <150> PRIOR APPLICATION NUMBER: US/09/939,226
 12 <151> PRIOR FILING DATE: 2001-08-24
 13 <150> PRIOR APPLICATION NUMBER: US 08/833,752
 14 <151> PRIOR FILING DATE: 1997-04-09
 15 <150> PRIOR APPLICATION NUMBER: US 09/626,939
 16 <151> PRIOR FILING DATE: 2000-07-27
 17 <150> PRIOR APPLICATION NUMBER: US 08/810,028
 18 <151> PRIOR FILING DATE: 1997-03-04
 19 <160> NUMBER OF SEQ ID NOS: 18
 20 <170> SOFTWARE: PatentIn version 3.1
 22 <210> SEQ ID NO: 1
 23 <211> LENGTH: 792
 24 <212> TYPE: DNA
 25 <213> ORGANISM: Homo sapiens
 26 <400> SEQUENCE: 1
 27       gaattccccc aacagagcca agctctccat ctagtggaca gggaagctag cagcaaacct      60
 28       tcccttcact acaaaacttc attgcttggc caaaaagaga gttaattcaa tgtagacatc     120
 29       tatgtaggca attaaaaacc tattgatgta taaaacagtt tgcattcatg gagggcaact     180
 30       aaatacattc taggacttta taaaagatca ctttttattt atgcacaggg tggaacaaga     240
 31       tggattatca agtgtcaagt ccaatctatg acatcaatta ttatacatcg gagccctgcc     300
 32       aaaaaatcaa tgtgaagcaa atcgcagccc gcctcctgcc tccgctctac tcactggtgt     360
 33       tcatctttgg ttttgtgggc aacatgctgg tcatcctcat cctgataaac tgcaaaaggc     420
 34       tgaagagcat gactgacatc tacctgctca acctggccat ctctgacctg ttttttcttc     480
 35       ttactgtccc cttctgggct cactatgctg ccgcccagtg ggactttgga aatacaatgt     540
 36       gtcaactctt gacagggctc tattttatag gcttcttctc tggaatcttc ttcatcatcc     600
 37       tcctgacaat cgataggtac ctggctgtcg tccatgctgt gtttgcttta aaagccagga     660
 38       cggtcacctt tggggtggtg acaagtgtga tcacttgggt ggtggctgtg tttgcgtctc     720
 39       tcccaggaat catctttacc agatctcaaa aagaaggtct tcattacacc tgcagctctc     780
 40       attttccata ca                                                         792
 42 <210> SEQ ID NO: 2
 43 <211> LENGTH: 1477
 44 <212> TYPE: DNA
 45 <213> ORGANISM: Homo sapiens
```

RAW SEQUENCE LISTING   DATE: 04/20/2005
PATENT APPLICATION: US/09/938,703B   TIME: 12:35:41

Input Set : **N:\Crf3\RULE60\09938703B.raw.txt**
Output Set: **N:\CRF4\04202005\I938703B.raw**

```
46 <220> FEATURE:
47 <221> NAME/KEY: misc_feature
48 <222> LOCATION: (1377)..(1377)
49 <223> OTHER INFORMATION: Any nucleotide
50 <220> FEATURE:
51 <221> NAME/KEY: misc_feature
52 <222> LOCATION: (1384)..(1385)
53 <223> OTHER INFORMATION: Any nucleotide
54 <400> SEQUENCE: 2
55      gaattccccc aacagagcca agctctccat ctagtggaca gggaagctag cagcaaacct      60
56      tcccttcact acaaaacttc attgcttggc caaaaagaga gttaattcaa tgtagacatc     120
57      tatgtaggca attaaaaacc tattgatgta taaaacagtt tgcattcatg gagggcaact     180
58      aaatacattc taggacttta taaaagatca ctttttattt atgcacaggg tggaacaaga     240
59      tggattatca agtgtcaagt ccaatctatg acatcaatta ttatacatcg gagccctgcc     300
60      aaaaaatcaa tgtgaagcaa atcgcagccc gcctcctgcc tccgctctac tcactggtgt     360
61      tcatctttgg ttttgtgggc aacatgctgg tcatcctcat cctgataaac tgcaaaaggc     420
62      tgaagagcat gactgacatc tacctgctca acctggccat ctctgacctg tttttccttc     480
63      ttactgtccc cttctgggct cactatgctg ccgcccagtg ggactttgga aatacaatgt     540
64      gtcaactctt gacagggctc tattttatag gcttcttctc tggaatcttc ttcatcatcc     600
65      tcctgacaat cgataggtac ctggctgtcg tccatgctgt gtttgcttta aaagccagga     660
66      cggtcacctt tggggtggtg acaagtgtga tcacttgggt ggtggctgtg tttgcgtctc     720
67      tcccaggaat catctttacc agatctcaaa aagaaggtct tcattacacc tgcagctctc     780
68      attttccata cagtcagtat caattctgga agaatttcca gacattaaag atagtcatct     840
69      tggggctggt cctgccgctg cttgtcatgg tcatctgcta ctcgggaatc ctaaaaactc     900
70      tgcttcggtg tcgaaatgag aagaagaggc acagggctgt gaggcttatc ttcaccatca     960
71      tgattgttta ttttctcttc tgggctccct acaacattgt ccttctcctg aacaccttcc    1020
72      aggaattctt tggcctgaat aattgcagta gctctaacag gttggaccaa gctatgcagg    1080
73      tgacagagac tcttgggatg acgcactgct gcatcaaccc catcatctat gcctttgtcg    1140
74      gggagaagtt cagaaactac ctcttagtct tcttccaaaa gcacattgcc aaacgcttct    1200
75      gcaaatgctg ttctattttc cagcaagagg ctcccgagcg agcaagctca gtttacaccc    1260
76      gatccactgg ggagcaggaa atatctgtgg gcttgtgaca cggactcaag tgggctggtg    1320
```
W--> 77   acccagtcag agttgtgcac atggcttagt tttcatacac agcctgggct gggggtnggt   1380
```
78      tggnngaggt ctttttttaaa aggaagttac tgttatagag ggtctaagat tcatccattt    1440
79      atttggcatc tgtttaaagt agattagatc cgaattc                              1477
81 <210> SEQ ID NO: 3
82 <211> LENGTH: 1442
83 <212> TYPE: DNA
84 <213> ORGANISM: Homo sapiens
85 <400> SEQUENCE: 3
86      gaattccccc aacagagcca agctctccat ctagtggaca gggaagctag cagcaaacct      60
87      tcccttcact acaaaacttc attgcttggc caaaaagaga gttaattcaa tgtagacatc     120
88      tatgtaggca attaaaaacc tattgatgta taaaacagtt tgcattcatg gagggcaact     180
89      aaatacattc taggacttta taaaagatca ctttttattt atgcacaggg tggaacaaga     240
90      tggattatca agtgtcaagt ccaatctatg acatcaatta ttatacatcg gagccctgcc     300
91      aaaaaatcaa tgtgaagcaa atcgcagccc gcctcctgcc tccgctctac tcactggtgt     360
92      tcatctttgg ttttgtgggc aacatgctgg tcatcctcat cctgataaac tgcaaaaggc     420
93      tgaagagcat gactgacatc tacctgctca acctggccat ctctgacctg tttttccttc     480
94      ttactgtccc cttctgggct cactatgctg ccgcccagtg ggactttgga aatacaatgt     540
95      gtcaactctt gacagggctc tattttatag gcttcttctc tggaatcttc ttcatcatcc     600
```

```
 96        tcctgacaat cgataggtac ctggctgtcg tccatgctgt gtttgcttta aaagccagga    660
 97        cggtcacctt tggggtggtg acaagtgtga tcacttgggt ggtggctgtg tttgcgtctc    720
 98        tcccaggaat catctttacc agatctcaaa aagaaggtct tcattacacc tgcagctctc    780
 99        attttccata cattaaagat agtcatcttg gggctggtcc tgccgctgct tgtcatggtc    840
100        atctgctact cgggaatcct aaaaactctg cttcggtgtc gaaatgagaa gaagaggcac    900
101        agggctgtga ggcttatctt caccatcatg attgtttatt ttctcttctg ggctccctac    960
102        aacattgtcc ttctcctgaa caccttccag gaattctttg cctgaataa ttgcagtagc   1020
103        tctaacaggt tggaccaagc tatgcaggtg acagagactc ttgggatgac gcactgctgc   1080
104        atcaacccca tcatctatgc ctttgtcggg gagaagttca gaaactacct cttagtcttc   1140
105        ttccaaaagc acattgccaa acgcttctgc aaatgctgtt ctattttcca gcaagaggct   1200
106        cccgagcgag caagctcagt ttacacccga tccactgggg agcaggaaat atctgtgggc   1260
107        ttgtgacacg gactcaagtg ggctggtgac ccagtcagag ttgtgcacat ggcttagttt   1320
108        tcatacacag cctgggctgg gggtggttgg gaggtctttt ttaaaaggaa gttactgtta   1380
109        tagagggtct aagattcatc catttatttg gcatctgttt aaagtagatt agatccgaat   1440
110        tc                                                                   1442
112 <210> SEQ ID NO: 4
113 <211> LENGTH: 184
114 <212> TYPE: PRT
115 <213> ORGANISM: Homo sapiens
116 <400> SEQUENCE: 4
117        Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr
118        1               5                   10                  15
119        Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu
120                        20                  25                  30
121        Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn
122                    35                  40                  45
123        Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met
124            50                  ·  55                  60
125        Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu
126        65                  70                  75                  80
127        Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Ala Gln Trp Asp Phe
128                        85                  90                  95
129        Gly Asn Thr Met Cys Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe
130                        100                 105                 110
131        Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu
132                    115                 120                 125
133        Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe
134            130                 135                 140
135        Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser
136        145                 150                 155                 160
137        Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr
138                        165                 170                 175
139        Thr Cys Ser Ser His Phe Pro Tyr
140                        180
142 <210> SEQ ID NO: 5
143 <211> LENGTH: 352
144 <212> TYPE: PRT
145 <213> ORGANISM: Homo sapiens
146 <400> SEQUENCE: 5
```

```
147      Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr
148      1               5                   10                  15
149      Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu
150                  20                  25                  30
151      Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn
152                  35                  40                  45
153      Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met
154          50                  55                  60
155      Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu
156      65                  70                  75                  80
157      Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Ala Gln Trp Asp Phe
158                  85                  90                  95
159      Gly Asn Thr Met Cys Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe
160                  100                 105                 110
161      Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu
162                  115                 120                 125
163      Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe
164          130                 135                 140
165      Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser
166      145                 150                 155                 160
167      Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr
168                  165                 170                 175
169      Thr Cys Ser Ser His Phe Pro Tyr Ser Gln Tyr Gln Phe Trp Lys Asn
170                  180                 185                 190
171      Phe Gln Thr Leu Lys Ile Val Ile Leu Gly Leu Val Leu Pro Leu Leu
172                  195                 200                 205
173      Val Met Val Ile Cys Tyr Ser Gly Ile Leu Lys Thr Leu Leu Arg Cys
174          210                 215                 220
175      Arg Asn Glu Lys Lys Arg His Arg Ala Val Arg Leu Ile Phe Thr Ile
176      225                 230                 235                 240
177      Met Ile Val Tyr Phe Leu Phe Trp Ala Pro Tyr Asn Ile Val Leu Leu
178                  245                 250                 255
179      Leu Asn Thr Phe Gln Glu Phe Phe Gly Leu Asn Asn Cys Ser Ser Ser
180                  260                 265                 270
181      Asn Arg Leu Asp Gln Ala Met Gln Val Thr Glu Thr Leu Gly Met Thr
182                  275                 280                 285
183      His Cys Cys Ile Asn Pro Ile Ile Tyr Ala Phe Val Gly Glu Lys Phe
184          290                 295                 300
185      Arg Asn Tyr Leu Leu Val Phe Phe Gln Lys His Ile Ala Lys Arg Phe
186      305                 310                 315                 320
187      Cys Lys Cys Cys Ser Ile Phe Gln Gln Glu Ala Pro Glu Arg Ala Ser
188                  325                 330                 335
189      Ser Val Tyr Thr Arg Ser Thr Gly Glu Gln Glu Ile Ser Val Gly Leu
190                  340                 345                 350
192 <210> SEQ ID NO: 6
193 <211> LENGTH: 215
194 <212> TYPE: PRT
195 <213> ORGANISM: Homo sapiens
196 <400> SEQUENCE: 6
```

```
197        Met Asp Tyr Gln Val Ser Ser Pro Ile Tyr Asp Ile Asn Tyr Tyr Thr
198        1               5                   10                  15
199        Ser Glu Pro Cys Gln Lys Ile Asn Val Lys Gln Ile Ala Ala Arg Leu
200                        20                  25                  30
201        Leu Pro Pro Leu Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn
202                35                  40                  45
203        Met Leu Val Ile Leu Ile Leu Ile Asn Cys Lys Arg Leu Lys Ser Met
204            50                  55                  60
205        Thr Asp Ile Tyr Leu Leu Asn Leu Ala Ile Ser Asp Leu Phe Phe Leu
206        65                  70                  75                  80
207        Leu Thr Val Pro Phe Trp Ala His Tyr Ala Ala Ala Gln Trp Asp Phe
208                        85                  90                  95
209        Gly Asn Thr Met Cys Gln Leu Leu Thr Gly Leu Tyr Phe Ile Gly Phe
210                        100                 105                 110
211        Phe Ser Gly Ile Phe Phe Ile Ile Leu Leu Thr Ile Asp Arg Tyr Leu
212                115                 120                 125
213        Ala Val Val His Ala Val Phe Ala Leu Lys Ala Arg Thr Val Thr Phe
214            130                 135                 140
215        Gly Val Val Thr Ser Val Ile Thr Trp Val Val Ala Val Phe Ala Ser
216        145                 150                 155                 160
217        Leu Pro Gly Ile Ile Phe Thr Arg Ser Gln Lys Glu Gly Leu His Tyr
218                        165                 170                 175
219        Thr Cys Ser Ser His Phe Pro Tyr Ile Lys Asp Ser His Leu Gly Ala
220                        180                 185                 190
221        Gly Pro Ala Ala Ala Cys His Gly His Leu Leu Leu Gly Asn Pro Lys
222                195                 200                 205
223        Asn Ser Ala Ser Val Ser Lys
224            210                 215
226 <210> SEQ ID NO: 7
227 <211> LENGTH: 360
228 <212> TYPE: PRT
229 <213> ORGANISM: Homo sapiens
230 <220> FEATURE:
231 <221> NAME/KEY: MISC_FEATURE
232 <222> LOCATION: (325)..(327)
233 <223> OTHER INFORMATION: Xaa = any amino acid
234 <400> SEQUENCE: 7
235        Met Leu Ser Thr Ser Arg Ser Arg Phe Ile Arg Asn Thr Asn Glu Ser
236        1               5                   10                  15
237        Gly Glu Glu Val Thr Thr Phe Phe Asp Tyr Asp Tyr Gly Ala Pro Cys
238                        20                  25                  30
239        His Lys Phe Asp Val Lys Gln Ile Gly Ala Gln Leu Leu Pro Pro Leu
240                35                  40                  45
241        Tyr Ser Leu Val Phe Ile Phe Gly Phe Val Gly Asn Met Leu Val Val
242            50                  55                  60
243        Leu Ile Leu Ile Asn Cys Lys Lys Leu Lys Cys Leu Thr Asp Ile Tyr
244        65                  70                  75                  80
245        Leu Leu Asn Leu Ala Ile Ser Asp Leu Leu Phe Ile Ile Thr Leu Pro
246                        85                  90                  95
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 04/20/2005
PATENT APPLICATION: US/09/938,703B   TIME: 12:35:42

Input Set : N:\Crf3\RULE60\09938703B.raw.txt
Output Set: N:\CRF4\04202005\I938703B.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:2; N Pos. 1377,1384,1385
Seq#:7; Xaa Pos. 325,326,327
Seq#:8; Xaa Pos. 231,232,233,333,334,335
Seq#:10; Xaa Pos. 145,146,147,321,322,323

Invalid Line Length:
The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 6
Seq#:14; Line(s) 494
Seq#:15; Line(s) 506

VERIFICATION SUMMARY   DATE: 04/20/2005
PATENT APPLICATION: US/09/938,703B   TIME: 12:35:42

Input Set : **N:\Crf3\RULE60\09938703B.raw.txt**
Output Set: **N:\CRF4\04202005\I938703B.raw**

L:9 M:270 C: Current Application Number differs, Wrong Format
L:77 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:1320
M:341 Repeated in SeqNo=2
L:275 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:320
L:323 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:224
M:341 Repeated in SeqNo=8
L:425 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:144
M:341 Repeated in SeqNo=10